Analyze Decentralized Personal Health Data using Solid, Digital Consent, and Federated Learning

Chang Sun, Ph.D.
Institute of Data Science, Maastricht University

Solid Symposium 2023 - Applications for Health Data Management with Solid
31 March 2023
Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data and repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA relevance)

OJ L 119, 4.5.2016, p. 1–88 (BG, ES, CS, DA, DE, ET, EL, EN, FR, GA, HR, IT, LV, LT, HU, MT, NL, PL, PT, RO, SK, SL, FI, SV)

In force: This act has been changed. Current consolidated version: 04/05/2016

A lot of work focused on analysing the data controlled by organizations, but what about using data directly controlled by individuals?
Citizens would like to have more access and control over the data they generate.

EU should make major investments in technologies and infrastructures that enhance data access and use, while giving us as well as public and private organisations full control over the data they generate.

Law and technology should enable us to make available their data for the public interest without any direct reward.

We would be willing to make such data available, especially for health-related research and for aspects relating to the locality we live in (e.g. mobility, environment).

There are no sufficient tools and mechanisms for us to ‘donate’ our data.

A key challenge

“Individuals value the high level of protection granted by the GDPR and ePrivacy legislation. However, they suffer from the absence of technical tools and standards that make the exercise of their rights simple and not overly burdensome.”

-- European Committee <An European Strategy for Data> 2020

Solutions:

provide individuals the tools to decide at a granular level what is done with their data using consent management, personal information management, decentralised networking, etc, in order to give greater oversight, transparency, control over their personal data.
citizen-centric DATA platform (TIDAL)

Based on Solid (SOfcial Linked Data) to store and structure personal data, and give individual control access to their fine-grained personal data.
Based on Solid (SOcial Linked Data) to store and structure personal data, and give individual control access to their fine-grained personal data.

Integrating Data Privacy Vocabulary to structure personal data requests as digital consents.

Data Privacy Vocabulary: https://w3c.github.io/dpv/dpv/
Based on Solid (SOcial LInked Data) to store and structure personal data, and give individual control access to their fine-grained personal data. Integrating Data Privacy Vocabulary to structure personal data requests as digital consents. Formulating personal data and data request into RDF format with integration of vocabulary services and standards.

Listing 4: An example of the RDF data file in a participant’s Solid pod.

```
@prefix : <https://exampleparticipant.solidprovider.com/profile/card#>.
@prefix SNOMEDCT: <http://purl.bioontology.org/ontology/SNOMEDCT/>.

:me a SNOMEDCT:116154003; # Patient
  SNOMEDCT:397669002 "27"^^xsd:int; # Age
  SNOMEDCT:50373000 "165"^^xsd:int; # Height
  SNOMEDCT:726527001 "55"^^xsd:int; # Weight
  SNOMEDCT:263495000 SNOMEDCT:248152002; # Gender, Female
  SNOMEDCT:271649006 "110"; # Systolic blood pressure
  SNOMEDCT:271650006 "90"; # Diastolic blood pressure
  SNOMEDCT:405751000 SNOMEDCT:44054006. # Type 2 diabetes
```
Based on Solid (SOcial Linked Data) to store and structure personal data, and give individual control access to their fine-grained personal data.

Integrating Data Privacy Vocabulary to structure personal data requests as digital consents to meet the requirements of GDPR.

Formulating personal data and data request into RDF format with integration of vocabulary services and standards.

Integrating federated learning (Personal Health Train) using the parameters promised in the data request and only the results are sent to the researchers.
TIDAL framework - Data request and consent

1. Fill in
2. Store
3. View & approve
4. Register WebID to the request
4. Granted access to Trusted party

Please Note: This request form is structured using the Data Privacy Vocabulary (DPV). DPV provides terms (classes and properties) to describe and represent information related to processing of personal data based on established requirements such as GDPR.

Purpose of your research

- Research and Development

Description of your purpose:
Learn association between diabetes status and dietary pattern

Personal data categories
- Medical Health [Special] (Physical Health, Mental Health, DNA Code, Disability, Health History)
- Demographic (Physical Trait, Income Bracket, Geographic)

Data elements (URI)
- Diagnosis

Consent Duration (Days)
- 90

Number of instances (minimal)
- 100

Data Processing Category
- CRISP: Diagnosis

Analysis Model
- Linear Regression

Consequences of data processing and impact of your research:
Help diabetes patients understand the impact of their diet pattern
TIDAL framework - Data request and consent

@prefix : <http://exampleresearcher.solidprovider.com/public/request.ttl#>.
@prefix schema: <https://schema.org/>.
@prefix exre: <http://exampleresearcher.solidprovider.com/profile/card#>.
@prefix dpv: <http://w3id.org/dpv#>.
@prefix dpvpd: <http://w3id.org/dpv/dpv-pd#>.
@prefix dpvtech: <http://w3id.org/dpv/dpv-tech#>.
@prefix SNOMEDCT: <http://purl.bioontology.org/ontology/SNOMEDCT/>.

:a schema:AskAction, dpv:PersonalDataHandling;
  dpv:hasLegalBasis dpv:Consent;
  rdfs:label "Learn association between diabetes status and dietary pattern";
  schema:collectionSize 32;
  schema:creator exre:me;
  schema:DataFeedItem SNOMEDCT:10396001, SNOMEDCT:230125005, SNOMEDCT:56718006, SNOMEDCT:73211009;
  schema:dateCreated "2021-01-18T00:00:00Z"^^XML:dateTime;
  dpvtech:isImplementedUsingTechnology dpvtech:LinearRegression;
  dpv:hasImpact "Help diabetes patients understand the impact of their diet pattern";
  dpv:hasContext SNOMEDCT;
  dpv:hasDataController exre:me;
  dpv:hasExpiryTime "2021-12-31T00:00:00Z"^^XML:dateTime;
  dpvpd:hasPersonalData dpvpd:Health;
  dpv:hasProcessing dpv:Analyse;
  dpv:hasPurpose dpv:ResearchAndDevelopment.
TIDAL framework - Data request and consent

All ongoing data requests

Viewing
View the published data requests from researchers.

Period
Enter the specific end date for using your data.

Respond
Click decline or approve to the data request.

CommercialInterest

Researcher A
Institute A
Class of purpose: CommercialInterest
Purpose: How XXX health app can help patient with obesity to have healthier lifestyle.
Personal Data Category: HealthRecord
Data Processing Category: Analyse
Requested data:
- Weight change
- Running
- Body mass index
End date: 2021-08-15T00:00:00Z
Instances: 300
Analysis: Linear Regression
Withdrawal Date: dd/mm/yyyy
Data Recipient:

Researcher B
Institute B
Class of purpose: ResearchAndDevelopment
Purpose: Learn association between diabetes status and dietary pattern
Personal Data Category: Health
Data Processing Category: Analyse
Requested data:
- Hyperlipoproteinemia diet
- Diabetes mellitus
End date: 2021-08-01T00:00:00Z
Instances: 32
Analysis: Linear Regression
Consequence of data process: Help diabetes patients understand the impact of their diet pattern
Requested data is not detected in your pod!

Researcher C
Institute B
Class of purpose: Security
Purpose: Improve the security of health data storage and sharing
Personal Data Category: HealthRecord
Data Processing Category: Profiling
PseudoAnonymise
Requested data:
- Blood pressure
- Heart rate
- Plasma glucose
End date: 2022-01-01T00:00:00Z
Instances: 265
Analysis: Linear Regression
Requested data is not detected in your pod!
Listing 2: An example of generated participation statements in a RDF format in the participant’s Solid pod.

```json
@prefix : <http://exampleParticipant.solidProvider.com/private/participation#>.  
@prefix part: <https://exampleparticipant.solidprovider.com/profile/card#>.  
@prefix req: <https://exampleresearcher.solidprovider.com/public/request.ttl#>.  
@prefix extp: <http://exampleretrustedparty.solidprovider.net/profile/card#> 
@prefix app: <https://examplesolidapp.com/>

:16197041266295299657542155198  
a schema:JoinAction, dpv:Consent;  
schema:dateCreated "2021-02-18T00:00:00Z"^^XML:dateTime;  
dpv:DataSubject part:me;  
dpv:hasConsentNotice req:161964062096710764675982245664;  
dpv:hasExpire schema:false;  
dpv:hasExpireTime "2021-12-31T00:00:00Z"^^XML:dateTime;  
dpv:hasProvisionMethod app:participate;  
dpv:hasProvisionTime "2021-02-18T00:00:00Z"^^XML:dateTime;  
dpv:hasWithdrawalTime "2021-09-18T00:00:00Z"^^XML:dateTime;  
dpv:hasRecipient extp:me.  
```
TIDAL framework - Data Analysis
TIDAL framework - Data Analysis

Listing 3: An example of generated trigger message (Activate Action) sent by the researcher.

@prefix : <http://examplertrustedparty.solidprovider.net/inbox/triggermessage#>.
@prefix req: <http://exampleresearcher.solidprovider.com/public/request.ttl#>.
@prefix exre: <http://exampleresearcher.solidprovider.com/profile/card#>.

:160622932739325095672093710975
  schema:actionStatus schema:ActivateAction;
  schema:creator exre:me;
  schema:dateCreated "2021-04-20T09:37:57.499Z"^^XML:dateTime;
  schema:target req:16196406209671076675982245664.
TIDAL framework - Data Analysis
Discussion

- Individuals data will be read only when the analysis gets executed.
- The analysis can only be triggered by the data requester under conditions (have enough participants, in the valid period, for the indicated purpose, etc)
- Individuals can revoke the permission anytime.
- Individuals can change the data values at any moment (even after giving researchers permission, when the experiments are being conducted)
- (Future work) Individuals will be able to indicate their preference to which data requests they want to receive.
- and more…
Paper: ciTizen-centric DAta pLatform (TIDAL): Sharing Distributed Personal Data in a Privacy-Preserving Manner for Health Research

Open Source tool and code: https://github.com/sunchang0124/TIDAL

Web application is live →

Demo Youtube Video:

Contact:
- chang.sun@maastrichtuniversity.nl
- www.linkedin.com/in/chang-sun-maastricht/